UNITED STATES PATENT AND TRADEMARK OFFICE

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/664,450	09/20/2003	Yurdaer Nezihi Doganata	CHA920030010US1	4523

7590   04/17/2008

James R. Murray
207 Dogwood Court
Poughkeepsie, NY 12601

EXAMINER
DWIVEDI, MAHESH H

ART UNIT	PAPER NUMBER
2168	

MAIL DATE	DELIVERY MODE
04/17/2008	PAPER

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

	Application No.	Applicant(s)
Office Action Summary	10/664,450	DOGANATA ET AL.
	Examiner	Art Unit	
	MAHESH H. DWIVEDI	2168	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE *3* MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on *07 January 2008*.
2a)☐ This action is **FINAL**.　　　　2b)☒ This action is non-final.
3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *1,4-9 and 12-20* is/are pending in the application.
　　　4a) Of the above claim(s) _____ is/are withdrawn from consideration.
5)☐ Claim(s) _____ is/are allowed.
6)☒ Claim(s) *1,4-9 and 12-20* is/are rejected.
7)☐ Claim(s) _____ is/are objected to.
8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.
10)☒ The drawing(s) filed on *16 March 2007* is/are: a)☐ accepted or b)☐ objected to by the Examiner.
　　　Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).
　　　Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
　　　a)☐ All　b)☐ Some * c)☐ None of:
　　　　　1.☐ Certified copies of the priority documents have been received.
　　　　　2.☐ Certified copies of the priority documents have been received in Application No. _____.
　　　　　3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).
　　　* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☐ Information Disclosure Statement(s) (PTO/SB/08)
　　　Paper No(s)/Mail Date _____.
4) ☐ Interview Summary (PTO-413)
　　　Paper No(s)/Mail Date. _____ .
5) ☐ Notice of Informal Patent Application
6) ☐ Other: _____.

DETAILED ACTION

Continued Examination Under 37 CFR 1.114

1.   A request for continued examination under 37 CFR 1.114, including the fee set forth in 37 CFR 1.17(e), was filed in this application after final rejection. Since this application is eligible for continued examination under 37 CFR 1.114, and the fee set forth in 37 CFR 1.17(e) has been timely paid, the finality of the previous Office action has been withdrawn pursuant to 37 CFR 1.114. Applicant's submission filed on 9/5/2006 has been entered.

Response to Amendment

2.   Receipt of Applicant's Amendment, filed on 01/07/2008, is acknowledged. The amendment includes the amending of the specification, the amending of claims 1, 8-9, and 17-18, and the cancellation of claims 2-3, and 10-11.

Information Disclosure Statement

3.   The information disclosure statement filed 03/16/2007 fails to comply with 37 CFR 1.98(a)(2), which requires a legible copy of each cited foreign patent document; each non-patent literature publication or that portion which caused it to be listed; and all other information or that portion which caused it to be listed. It has been placed in the application file, but the information referred to therein has not been considered. The examiner specifically points to references 1-2, 5, and 7 as patent documents which are not published, and as a result should be moved to the non-patent literature section, accompanied with copies of those references.

The information disclosure statement (IDS) submitted on 03/16/2007 has been received, entered into the record, and considered. The submission is in compliance with the provisions of 37 CFR 1.97. Accordingly, the information disclosure statement is being considered by the examiner.

Claim Rejections - 35 USC § 112

4.   The following is a quotation of the second paragraph of 35 U.S.C. 112:

> The specification shall conclude with one or more claims particularly pointing out and distinctly claiming the subject matter which the applicant regards as his invention.

5.   Claim 17 recites the limitation "looks through **the** search system log" in page 12.
There is insufficient antecedent basis for this limitation in the claim.

Claims 18-20 are rejected for incorporating the deficiencies of claim 17.

Claim Objections

6.   Claim 1 is objected to because of the following informalities: The phrase "**An**
self-enhancing search system comprising:" should be changed to "**A** self-enhancing
search system comprising:". Appropriate correction is required.

Claims 4-8 are objected to for incorporating the deficiencies of independent claim
1.

Claim Rejections - 35 USC § 102

7.   The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that
form the basis for the rejections under this section made in this Office action:

> A person shall be entitled to a patent unless –
> (b) the invention was patented or described in a printed publication in this or a foreign country or in public
> use or on sale in this country, more than one year prior to the date of application for patent in the United
> States.

8.   Claims 1, 4-9. and 12-20 rejected under 35 U.S.C. 102(b) as being anticipated by
Doganata et al. (Article entitled "DBlue: An Advanced Enterprise Information Search
and Delivery System", published on January 01, 2000).

9.   Regarding claim 1, **Doganata** teaches a search system comprising:

A) a search system analysis system that periodically looks through a log of search
queries of the search system and identifies, for analyzing, unsatisfactory, for analyzing,
unsatisfactory customer search queries that do not bring satisfactory results from a
database being searched by the customer (Pages 4-5);

B) a search query analyzer using one or more of glossary terms, synonyms, known
typographical errors and translated words to provide alternative query terms to original
search query terms in the unsatisfactory search queries (Pages 4-5);

C) a relevant document finder based on enhanced queries including the alternative
query terms to locate relevant documents in the searched database not found when the
unsatisfactory identified search queries were used (Pages 4-5); and

D) <u>embedding in located relevant documents not found by the unsuccessful search</u> <u>queries those of the original search query terms not contained in those relevant</u> <u>documents</u> (Pages 4-5).

The examiner further notes that **Doganata** teaches **"a search system analysis system that periodically looks through a log <u>of</u> search queries <u>of</u> the search system and identifies, for analyzing, unsatisfactory, for analyzing, unsatisfactory <u>customer</u> search queries that do not bring satisfactory results <u>from a database</u> <u>being searched by the customer</u>"** as "The search results contain the documents that are indexed against the query terms and scored based on certain statistical criteria. In many real-life situations, the relevant documents can't be found or may not appear at the top of the search results because they are scored low or they don't contain the terms exactly as in the query. This is common when users choose variations of the query terms, including inflections, misspellings, abbreviations, and so on. To improve the user experience, dBlue uses text analysis tools developed by IBM Research to enhance the contents of documents. This process is started by extracting terms from a large collection of documents in the IBM technical support domain to create a domain-specific glossary. The terms in the glossary can consist of canonical form, variant form (inflection, abbreviation, misspelling, etc.), synonym, term definition, statistical data, and other information. This initial glossary is enhanced by eliminating irrelevant terms and reranking terms using special dictionaries and algorithms" (Pages 4-5) and "Based on the glossary, the important keywords in each document are extracted and ranked, and their related glossary terms (variants, synonyms, etc.) are used to enrich the content of the document. The content enrichment is used to create keyword metatags for biased indexing, expand the query terms to include related terms, and enable search for related documents. To improve the user's search experience, keywords are displayed in the search results and navigating through keywords is possible" (Pages 4-5). The examiner further notes that **Doganata** teaches **"a search query analyzer using one or more of glossary terms, synonyms, known typographical errors and translated words to provide alternative query terms to original search <u>query</u> terms in the unsatisfactory search queries"** as "The search results contain the documents that are

indexed against the query terms and scored based on certain statistical criteria. In many real-life situations, the relevant documents can't be found or may not appear at the top of the search results because they are scored low or they don't contain the terms exactly as in the query. This is common when users choose variations of the query terms, including inflections, misspellings, abbreviations, and so on. To improve the user experience, dBlue uses text analysis tools developed by IBM Research to enhance the contents of documents. This process is started by extracting terms from a large collection of documents in the IBM technical support domain to create a domain-specific glossary. The terms in the glossary can consist of canonical form, variant form (inflection, abbreviation, misspelling, etc.), synonym, term definition, statistical data, and other information. This initial glossary is enhanced by eliminating irrelevant terms and reranking terms using special dictionaries and algorithms" (Pages 4-5) and "Based on the glossary, the important keywords in each document are extracted and ranked, and their related glossary terms (variants, synonyms, etc.) are used to enrich the content of the document. The content enrichment is used to create keyword metatags for biased indexing, expand the query terms to include related terms, and enable search for related documents. To improve the user's search experience, keywords are displayed in the search results and navigating through keywords is possible" (Pages 4-5). The examiner further notes that **Doganata** teaches **"a relevant document finder based on enhanced queries including the alternative query terms to locate relevant documents <u>in the searched database</u> not found when the unsatisfactory identified search queries were used"** as "The search results contain the documents that are indexed against the query terms and scored based on certain statistical criteria. In many real-life situations, the relevant documents can't be found or may not appear at the top of the search results because they are scored low or they don't contain the terms exactly as in the query. This is common when users choose variations of the query terms, including inflections, misspellings, abbreviations, and so on. To improve the user experience, dBlue uses text analysis tools developed by IBM Research to enhance the contents of documents. This process is started by extracting terms from a large collection of documents in the IBM technical support domain to create a domain-specific

glossary. The terms in the glossary can consist of canonical form, variant form (inflection, abbreviation, misspelling, etc.), synonym, term definition, statistical data, and other information. This initial glossary is enhanced by eliminating irrelevant terms and reranking terms using special dictionaries and algorithms" (Pages 4-5) and "Based on the glossary, the important keywords in each document are extracted and ranked, and their related glossary terms (variants, synonyms, etc.) are used to enrich the content of the document. The content enrichment is used to create keyword metatags for biased indexing, expand the query terms to include related terms, and enable search for related documents. To improve the user's search experience, keywords are displayed in the search results and navigating through keywords is possible" (Pages 4-5). The examiner further notes that **Doganata** teaches **"embedding in located relevant documents not found by the unsuccessful search queries those of the original search query terms not contained in those relevant documents"** as "The search results contain the documents that are indexed against the query terms and scored based on certain statistical criteria. In many real-life situations, the relevant documents can't be found or may not appear at the top of the search results because they are scored low or they don't contain the terms exactly as in the query. This is common when users choose variations of the query terms, including inflections, misspellings, abbreviations, and so on. To improve the user experience, dBlue uses text analysis tools developed by IBM Research to enhance the contents of documents. This process is started by extracting terms from a large collection of documents in the IBM technical support domain to create a domain-specific glossary. The terms in the glossary can consist of canonical form, variant form (inflection, abbreviation, misspelling, etc.), synonym, term definition, statistical data, and other information. This initial glossary is enhanced by eliminating irrelevant terms and reranking terms using special dictionaries and algorithms" (Pages 4-5) and "Based on the glossary, the important keywords in each document are extracted and ranked, and their related glossary terms (variants, synonyms, etc.) are used to enrich the content of the document. The content enrichment is used to create keyword metatags for biased indexing, expand the query terms to include related terms, and enable search for related documents. To improve

the user's search experience, keywords are displayed in the search results and navigating through keywords is possible" (Pages 4-5).

Regarding claim 4, **Doganata** further teaches a computer program comprising:
A) associating enhanced queries with the unsatisfactory search queries in the search system log for use with further queries (Pages 4-5).

The examiner notes that **Doganata** teaches **"associating enhanced queries with the unsatisfactory search queries in the search system log for use with further queries"** as "The search results contain the documents that are indexed against the query terms and scored based on certain statistical criteria. In many real-life situations, the relevant documents can't be found or may not appear at the top of the search results because they are scored low or they don't contain the terms exactly as in the query. This is common when users choose variations of the query terms, including inflections, misspellings, abbreviations, and so on. To improve the user experience, dBlue uses text analysis tools developed by IBM Research to enhance the contents of documents. This process is started by extracting terms from a large collection of documents in the IBM technical support domain to create a domain-specific glossary. The terms in the glossary can consist of canonical form, variant form (inflection, abbreviation, misspelling, etc.), synonym, term definition, statistical data, and other information. This initial glossary is enhanced by eliminating irrelevant terms and reranking terms using special dictionaries and algorithms" (Pages 4-5) and "Based on the glossary, the important keywords in each document are extracted and ranked, and their related glossary terms (variants, synonyms, etc.) are used to enrich the content of the document. The content enrichment is used to create keyword metatags for biased indexing, expand the query terms to include related terms, and enable search for related documents. To improve the user's search experience, keywords are displayed in the search results and navigating through keywords is possible" (Pages 4-5).

Regarding claim 5, **Doganata** further teaches a search system comprising:

A) including ranking the results of searches using both the unsatisfactory and the enhanced queries (Pages 4-5).

The examiner notes that **Doganata** teaches **"including ranking the results of searches using both the unsatisfactory and the enhanced queries"** as "The search results contain the documents that are indexed against the query terms and scored based on certain statistical criteria. In many real-life situations, the relevant documents can't be found or may not appear at the top of the search results because they are scored low or they don't contain the terms exactly as in the query. This is common when users choose variations of the query terms, including inflections, misspellings, abbreviations, and so on. To improve the user experience, dBlue uses text analysis tools developed by IBM Research to enhance the contents of documents. This process is started by extracting terms from a large collection of documents in the IBM technical support domain to create a domain-specific glossary. The terms in the glossary can consist of canonical form, variant form (inflection, abbreviation, misspelling, etc.), synonym, term definition, statistical data, and other information. This initial glossary is enhanced by eliminating irrelevant terms and reranking terms using special dictionaries and algorithms" (Pages 4-5) and "Based on the glossary, the important keywords in each document are extracted and ranked, and their related glossary terms (variants, synonyms, etc.) are used to enrich the content of the document. The content enrichment is used to create keyword metatags for biased indexing, expand the query terms to include related terms, and enable search for related documents. To improve the user's search experience, keywords are displayed in the search results and navigating through keywords is possible" (Pages 4-5).

Regarding claim 6, **Doganata** further teaches a search system comprising:
A) wherein the search query analyzer comprises a module including: a sub-module that identifies domain specific terms in a given query, using domain specific glossary (Pages 4-5);
B) a sub-module that finds synonyms and related terms for the identified terms, using domain specific thesaurus (Pages 4-5);

C) a sub-module that finds other statistically close terms (Pages 4-5); and

D) a sub-module that identifies relevant domain specific categories for the identified terms, using domain specific ontology (Pages 4-5).

The examiner notes that **Doganata** teaches **"wherein the search query analyzer comprises a module including: a sub-module that identifies domain specific terms in a given query, using domain specific glossary"** as "The search results contain the documents that are indexed against the query terms and scored based on certain statistical criteria. In many real-life situations, the relevant documents can't be found or may not appear at the top of the search results because they are scored low or they don't contain the terms exactly as in the query. This is common when users choose variations of the query terms, including inflections, misspellings, abbreviations, and so on. To improve the user experience, dBlue uses text analysis tools developed by IBM Research to enhance the contents of documents. This process is started by extracting terms from a large collection of documents in the IBM technical support domain to create a domain-specific glossary. The terms in the glossary can consist of canonical form, variant form (inflection, abbreviation, misspelling, etc.), synonym, term definition, statistical data, and other information. This initial glossary is enhanced by eliminating irrelevant terms and reranking terms using special dictionaries and algorithms" (Pages 4-5). The examiner further notes that **Doganata** teaches **"a sub-module that finds synonyms and related terms for the identified terms, using domain specific thesaurus"** as "The search results contain the documents that are indexed against the query terms and scored based on certain statistical criteria. In many real-life situations, the relevant documents can't be found or may not appear at the top of the search results because they are scored low or they don't contain the terms exactly as in the query. This is common when users choose variations of the query terms, including inflections, misspellings, abbreviations, and so on. To improve the user experience, dBlue uses text analysis tools developed by IBM Research to enhance the contents of documents. This process is started by extracting terms from a large collection of documents in the IBM technical support domain to create a domain-specific glossary. The terms in the glossary can consist of canonical form, variant form

(inflection, abbreviation, misspelling, etc.), synonym, term definition, statistical data, and other information. This initial glossary is enhanced by eliminating irrelevant terms and reranking terms using special dictionaries and algorithms" (Pages 4-5). The examiner further notes that **Doganata** teaches **"a sub-module that finds other statistically close terms"** as "The search results contain the documents that are indexed against the query terms and scored based on certain statistical criteria. In many real-life situations, the relevant documents can't be found or may not appear at the top of the search results because they are scored low or they don't contain the terms exactly as in the query. This is common when users choose variations of the query terms, including inflections, misspellings, abbreviations, and so on. To improve the user experience, dBlue uses text analysis tools developed by IBM Research to enhance the contents of documents. This process is started by extracting terms from a large collection of documents in the IBM technical support domain to create a domain-specific glossary. The terms in the glossary can consist of canonical form, variant form (inflection, abbreviation, misspelling, etc.), synonym, term definition, statistical data, and other information. This initial glossary is enhanced by eliminating irrelevant terms and reranking terms using special dictionaries and algorithms" (Pages 4-5). The examiner further notes that **Doganata** teaches **"a sub-module that identifies relevant domain specific categories for the identified terms, using domain specific ontology"** as "The search results contain the documents that are indexed against the query terms and scored based on certain statistical criteria. In many real-life situations, the relevant documents can't be found or may not appear at the top of the search results because they are scored low or they don't contain the terms exactly as in the query. This is common when users choose variations of the query terms, including inflections, misspellings, abbreviations, and so on. To improve the user experience, dBlue uses text analysis tools developed by IBM Research to enhance the contents of documents. This process is started by extracting terms from a large collection of documents in the IBM technical support domain to create a domain-specific glossary. The terms in the glossary can consist of canonical form, variant form (inflection, abbreviation, misspelling, etc.), synonym, term definition, statistical data, and other information. This

initial glossary is enhanced by eliminating irrelevant terms and reranking terms using special dictionaries and algorithms" (Pages 4-5).

Regarding claim 7, **Doganata** further teaches a search system comprising:
A) wherein the document finder comprises a module including the following sub-modules: a sub-module that finds documents in the identified categories, using an original textual index (Pages 4-5); and

B) a sub-module that filters the found documents to find additional relevant documents, based on the identified domain specific terms, synonyms, related terms, and statistically close terms (Pages 4-6).

The examiner notes that **Doganata** teaches **"wherein the document finder comprises a module including the following sub-modules: a sub-module that finds documents in the identified categories, using an original textual index"** as "The way the system is architected allows combining keyword search with navigational search. Based on a topic or a document type, users can narrow down search findings with a single click. This increases the chances of finding the requested information when the user query isn't specific enough to narrow down the search results on the first attempt. The categorized results are returned with the option of filtering the results based on IBM's product offerings and the document types." (Pages 4-5). The examiner further notes that **Doganata** teaches **"a sub-module that filters the found documents to find additional relevant documents, based on the identified domain specific terms, synonyms, related terms, and statistically close terms"** as "This increases the chances of finding the requested information when the user query isn't specific enough to narrow down the search results on the first attempt. The categorized results are returned with the option of filtering the results based on IBM's product offerings and the document types" (Pages 4-5), "The search results contain the documents that are indexed against the query terms and scored based on certain statistical criteria. In many real-life situations, the relevant documents can't be found or may not appear at the top of the search results because they are scored low or they don't contain the terms exactly as in the query. This is common when users choose

variations of the query terms, including inflections, misspellings, abbreviations, and so on. To improve the user experience, dBlue uses text analysis tools developed by IBM Research to enhance the contents of documents. This process is started by extracting terms from a large collection of documents in the IBM technical support domain to create a domain-specific glossary. The terms in the glossary can consist of canonical form, variant form (inflection, abbreviation, misspelling, etc.), synonym, term definition, statistical data, and other information. This initial glossary is enhanced by eliminating irrelevant terms and reranking terms using special dictionaries and algorithms" (Pages 4-5), and "In the near future, customers will be able to ask questions in natural language and the system won't require an exact match of words. In the near future, dBlue will also personalize searching so that once a user fills out a profile, responses will be filtered and ranked based on that profile. Multilanguage searches for documents written in Japanese, Chinese, and French will be supported by late 2002" (Page 6).

Regarding claim 8, **Doganata** further teaches a search system comprising:
A) including a linking meta-data enhancer with the following sub-modules: a sub-module that creates associations (links) between each found document and the given query (Pages 4-5); and
B) a sub-module that adds new doc-query links to the meta-data of the corresponding textual index entries (Pages 4-5).

The examiner notes that **Doganata** teaches "**including a linking meta-data enhancer with the following sub-modules: a sub-module that creates associations (links) between each found document and the given query**" as "The search results contain the documents that are indexed against the query terms and scored based on certain statistical criteria. In many real-life situations, the relevant documents can't be found or may not appear at the top of the search results because they are scored low or they don't contain the terms exactly as in the query. This is common when users choose variations of the query terms, including inflections, misspellings, abbreviations, and so on. To improve the user experience, dBlue uses text analysis tools developed by IBM Research to enhance the contents of documents. This

process is started by extracting terms from a large collection of documents in the IBM
technical support domain to create a domain-specific glossary. The terms in the
glossary can consist of canonical form, variant form (inflection, abbreviation,
misspelling, etc.), synonym, term definition, statistical data, and other information. This
initial glossary is enhanced by eliminating irrelevant terms and reranking terms using
special dictionaries and algorithms. The process of generating and enhancing the
glossary is semi-automatic, using glossary tools and the librarian. Figure 4 shows the
multiple components that compose the glossary of technical terms built for the dBlue
system" (Pages 4-5) and "Based on the glossary, the important keywords in each
document are extracted and ranked, and their related glossary terms (variants,
synonyms, etc.) are used to enrich the content of the document. The content
enrichment is used to create keyword metatags for biased indexing, expand the query
terms to include related terms, and enable search for related documents. To improve
the user's search experience, keywords are displayed in the search results and
navigating through keywords is possible" (Pages 4-5). The examiner further notes that
Doganata teaches **"a sub-module that adds new doc-query links to the meta-data
of the corresponding textual index entries"** as "The search results contain the
documents that are indexed against the query terms and scored based on certain
statistical criteria. In many real-life situations, the relevant documents can't be found or
may not appear at the top of the search results because they are scored low or they
don't contain the terms exactly as in the query. This is common when users choose
variations of the query terms, including inflections, misspellings, abbreviations, and so
on. To improve the user experience, dBlue uses text analysis tools developed by IBM
Research to enhance the contents of documents. This process is started by extracting
terms from a large collection of documents in the IBM technical support domain to
create a domain-specific glossary. The terms in the glossary can consist of canonical
form, variant form (inflection, abbreviation, misspelling, etc.), synonym, term definition,
statistical data, and other information. This initial glossary is enhanced by eliminating
irrelevant terms and reranking terms using special dictionaries and algorithms. The
process of generating and enhancing the glossary is semi-automatic, using glossary

tools and the librarian. Figure 4 shows the multiple components that compose the glossary of technical terms built for the dBlue system" (Pages 4-5) and "Based on the glossary, the important keywords in each document are extracted and ranked, and their related glossary terms (variants, synonyms, etc.) are used to enrich the content of the document. The content enrichment is used to create keyword metatags for biased indexing, expand the query terms to include related terms, and enable search for related documents. To improve the user's search experience, keywords are displayed in the search results and navigating through keywords is possible" (Pages 4-5).

Regarding claim 9, **Doganata** teaches a computer program comprising:

A) a search system analog system software module that periodically looks through a log for the search system and <u>selects</u> for analyzing unsuccessful <u>customer</u> search queries (Pages 4-5);

B) a search query analyzer software module using one or more of glossary terms, synonyms, known typographical errors and translated words to provide alternative query terms <u>to the terms used in the unsuccessful search queries</u> (Pages 4-5);

C) a relevant document finder software module <u>using</u> enhanced queries including the alternative query terms to locate relevant documents not found using said unsuccessful <u>customer</u> search queries (Pages 4-5); and

D) software for embedding search query terms of the unsuccessful queries in the documents located by the enhanced queries <u>and not found by the unsuccessful customer search so that the documents located by the enhanced queries will be found if the unsuccessful customer search queries are repeated</u> (Pages 4-5).

The examiner notes that **Doganata** teaches **"a search system analog system software module that periodically looks through a log for the search system and <u>selects</u> for analyzing unsuccessful <u>customer</u> search queries"** as "The search results contain the documents that are indexed against the query terms and scored based on certain statistical criteria. In many real-life situations, the relevant documents can't be found or may not appear at the top of the search results because they are scored low or they don't contain the terms exactly as in the query. This is common when

users choose variations of the query terms, including inflections, misspellings, abbreviations, and so on. To improve the user experience, dBlue uses text analysis tools developed by IBM Research to enhance the contents of documents. This process is started by extracting terms from a large collection of documents in the IBM technical support domain to create a domain-specific glossary. The terms in the glossary can consist of canonical form, variant form (inflection, abbreviation, misspelling, etc.), synonym, term definition, statistical data, and other information. This initial glossary is enhanced by eliminating irrelevant terms and reranking terms using special dictionaries and algorithms" (Pages 4-5) and "Based on the glossary, the important keywords in each document are extracted and ranked, and their related glossary terms (variants, synonyms, etc.) are used to enrich the content of the document. The content enrichment is used to create keyword metatags for biased indexing, expand the query terms to include related terms, and enable search for related documents. To improve the user's search experience, keywords are displayed in the search results and navigating through keywords is possible" (Pages 4-5). The examiner further notes that **Doganata** teaches **"a search query analyzer software module using one or more of glossary terms, synonyms, known typographical errors and translated words to provide alternative query terms <u>to the terms used in the unsuccessful search queries</u>"** as "The search results contain the documents that are indexed against the query terms and scored based on certain statistical criteria. In many real-life situations, the relevant documents can't be found or may not appear at the top of the search results because they are scored low or they don't contain the terms exactly as in the query. This is common when users choose variations of the query terms, including inflections, misspellings, abbreviations, and so on. To improve the user experience, dBlue uses text analysis tools developed by IBM Research to enhance the contents of documents. This process is started by extracting terms from a large collection of documents in the IBM technical support domain to create a domain-specific glossary. The terms in the glossary can consist of canonical form, variant form (inflection, abbreviation, misspelling, etc.), synonym, term definition, statistical data, and other information. This initial glossary is enhanced by eliminating irrelevant terms and

reranking terms using special dictionaries and algorithms" (Pages 4-5) and "Based on the glossary, the important keywords in each document are extracted and ranked, and their related glossary terms (variants, synonyms, etc.) are used to enrich the content of the document. The content enrichment is used to create keyword metatags for biased indexing, expand the query terms to include related terms, and enable search for related documents. To improve the user's search experience, keywords are displayed in the search results and navigating through keywords is possible" (Pages 4-5). The examiner further notes that **Doganata** teaches **"a relevant document finder software module <u>using</u> enhanced queries including the alternative query terms to locate relevant documents not found using said unsuccessful <u>customer</u> search queries"** as "The search results contain the documents that are indexed against the query terms and scored based on certain statistical criteria. In many real-life situations, the relevant documents can't be found or may not appear at the top of the search results because they are scored low or they don't contain the terms exactly as in the query. This is common when users choose variations of the query terms, including inflections, misspellings, abbreviations, and so on. To improve the user experience, dBlue uses text analysis tools developed by IBM Research to enhance the contents of documents. This process is started by extracting terms from a large collection of documents in the IBM technical support domain to create a domain-specific glossary. The terms in the glossary can consist of canonical form, variant form (inflection, abbreviation, misspelling, etc.), synonym, term definition, statistical data, and other information. This initial glossary is enhanced by eliminating irrelevant terms and reranking terms using special dictionaries and algorithms" (Pages 4-5) and "Based on the glossary, the important keywords in each document are extracted and ranked, and their related glossary terms (variants, synonyms, etc.) are used to enrich the content of the document. The content enrichment is used to create keyword metatags for biased indexing, expand the query terms to include related terms, and enable search for related documents. To improve the user's search experience, keywords are displayed in the search results and navigating through keywords is possible" (Pages 4-5). The examiner further notes that **Doganata** teaches **"software for embedding search**

query terms of the unsuccessful queries in the documents located by the enhanced queries <u>and not found by the unsuccessful customer search so that the documents located by the enhanced queries will be found if the unsuccessful customer search queries are repeated</u>" as "The search results contain the documents that are indexed against the query terms and scored based on certain statistical criteria. In many real-life situations, the relevant documents can't be found or may not appear at the top of the search results because they are scored low or they don't contain the terms exactly as in the query. This is common when users choose variations of the query terms, including inflections, misspellings, abbreviations, and so on. To improve the user experience, dBlue uses text analysis tools developed by IBM Research to enhance the contents of documents. This process is started by extracting terms from a large collection of documents in the IBM technical support domain to create a domain-specific glossary. The terms in the glossary can consist of canonical form, variant form (inflection, abbreviation, misspelling, etc.), synonym, term definition, statistical data, and other information. This initial glossary is enhanced by eliminating irrelevant terms and reranking terms using special dictionaries and algorithms" (Pages 4-5) and "Based on the glossary, the important keywords in each document are extracted and ranked, and their related glossary terms (variants, synonyms, etc.) are used to enrich the content of the document. The content enrichment is used to create keyword metatags for biased indexing, expand the query terms to include related terms, and enable search for related documents. To improve the user's search experience, keywords are displayed in the search results and navigating through keywords is possible" (Pages 4-5).

Regarding claim 12, **Doganata** further teaches a computer program comprising: A) software for providing associated enhanced queries with the unsatisfactory queries in the search system log for use in connection with further customer queries (Pages 4-5).

The examiner notes that **Doganata** teaches **"software for providing associated enhanced queries with the unsatisfactory queries in the search**

system log for use in connection with further customer queries" as "The search
results contain the documents that are indexed against the query terms and scored
based on certain statistical criteria. In many real-life situations, the relevant documents
can't be found or may not appear at the top of the search results because they are
scored low or they don't contain the terms exactly as in the query. This is common when
users choose variations of the query terms, including inflections, misspellings,
abbreviations, and so on. To improve the user experience, dBlue uses text analysis
tools developed by IBM Research to enhance the contents of documents. This process
is started by extracting terms from a large collection of documents in the IBM technical
support domain to create a domain-specific glossary. The terms in the glossary can
consist of canonical form, variant form (inflection, abbreviation, misspelling, etc.),
synonym, term definition, statistical data, and other information. This initial glossary is
enhanced by eliminating irrelevant terms and reranking terms using special dictionaries
and algorithms" (Pages 4-5) and "Based on the glossary, the important keywords in
each document are extracted and ranked, and their related glossary terms (variants,
synonyms, etc.) are used to enrich the content of the document. The content
enrichment is used to create keyword metatags for biased indexing, expand the query
terms to include related terms, and enable search for related documents. To improve
the user's search experience, keywords are displayed in the search results and
navigating through keywords is possible" (Pages 4-5).

        Regarding claim 13, **Doganata** further teaches a computer program comprising:
A) including software for ranking results of searches in order of their pertinency using
the enhances query terms as a ranking basis (Pages 4-5).

        The examiner notes that **Doganata** teaches **"including software for ranking
results of searches in order of their pertinency using the enhances query terms
as a ranking basis"** as "The search results contain the documents that are indexed
against the query terms and scored based on certain statistical criteria. In many real-life
situations, the relevant documents can't be found or may not appear at the top of the
search results because they are scored low or they don't contain the terms exactly as in

the query. This is common when users choose variations of the query terms, including inflections, misspellings, abbreviations, and so on. To improve the user experience, dBlue uses text analysis tools developed by IBM Research to enhance the contents of documents. This process is started by extracting terms from a large collection of documents in the IBM technical support domain to create a domain-specific glossary. The terms in the glossary can consist of canonical form, variant form (inflection, abbreviation, misspelling, etc.), synonym, term definition, statistical data, and other information. This initial glossary is enhanced by eliminating irrelevant terms and reranking terms using special dictionaries and algorithms" (Pages 4-5) and "Based on the glossary, the important keywords in each document are extracted and ranked, and their related glossary terms (variants, synonyms, etc.) are used to enrich the content of the document. The content enrichment is used to create keyword metatags for biased indexing, expand the query terms to include related terms, and enable search for related documents. To improve the user's search experience, keywords are displayed in the search results and navigating through keywords is possible" (Pages 4-5).

Regarding claim 14, **Doganata** further teaches a search system comprising:
A) wherein the search query analyzer software module comprises: a software module that identifies domain specific terms in a given query, using domain specific glossary (Pages 4-5);
B) a software sub-module that finds synonyms and related terms for the identified terms, using domain specific thesaurus (Pages 4-5);
C) a software sub-module that finds other statistically close terms (Pages 4-5); and
D) a software sub-module that identifies relevant domain specific categories for the identified terms, using domain specific ontology (Pages 4-5).

The examiner notes that **Doganata** teaches **"wherein the search query analyzer software module comprises: a software module that identifies domain specific terms in a given query, using domain specific glossary"** as "The search results contain the documents that are indexed against the query terms and scored based on certain statistical criteria. In many real-life situations, the relevant documents

can't be found or may not appear at the top of the search results because they are scored low or they don't contain the terms exactly as in the query. This is common when users choose variations of the query terms, including inflections, misspellings, abbreviations, and so on. To improve the user experience, dBlue uses text analysis tools developed by IBM Research to enhance the contents of documents. This process is started by extracting terms from a large collection of documents in the IBM technical support domain to create a domain-specific glossary. The terms in the glossary can consist of canonical form, variant form (inflection, abbreviation, misspelling, etc.), synonym, term definition, statistical data, and other information. This initial glossary is enhanced by eliminating irrelevant terms and reranking terms using special dictionaries and algorithms" (Pages 4-5). The examiner further notes that **Doganata** teaches **"a software sub-module that finds synonyms and related terms for the identified terms, using domain specific thesaurus"** as "The search results contain the documents that are indexed against the query terms and scored based on certain statistical criteria. In many real-life situations, the relevant documents can't be found or may not appear at the top of the search results because they are scored low or they don't contain the terms exactly as in the query. This is common when users choose variations of the query terms, including inflections, misspellings, abbreviations, and so on. To improve the user experience, dBlue uses text analysis tools developed by IBM Research to enhance the contents of documents. This process is started by extracting terms from a large collection of documents in the IBM technical support domain to create a domain-specific glossary. The terms in the glossary can consist of canonical form, variant form (inflection, abbreviation, misspelling, etc.), synonym, term definition, statistical data, and other information. This initial glossary is enhanced by eliminating irrelevant terms and reranking terms using special dictionaries and algorithms" (Pages 4-5). The examiner further notes that **Doganata** teaches **"a software sub-module that finds other statistically close terms"** as "The search results contain the documents that are indexed against the query terms and scored based on certain statistical criteria. In many real-life situations, the relevant documents can't be found or may not appear at the top of the search results because they are scored low or they don't contain the

terms exactly as in the query. This is common when users choose variations of the query terms, including inflections, misspellings, abbreviations, and so on. To improve the user experience, dBlue uses text analysis tools developed by IBM Research to enhance the contents of documents. This process is started by extracting terms from a large collection of documents in the IBM technical support domain to create a domain-specific glossary. The terms in the glossary can consist of canonical form, variant form (inflection, abbreviation, misspelling, etc.), synonym, term definition, statistical data, and other information. This initial glossary is enhanced by eliminating irrelevant terms and reranking terms using special dictionaries and algorithms" (Pages 4-5). The examiner further notes that **Doganata** teaches **"a software sub-module that identifies relevant domain specific categories for the identified terms, using domain specific ontology"** as "The search results contain the documents that are indexed against the query terms and scored based on certain statistical criteria. In many real-life situations, the relevant documents can't be found or may not appear at the top of the search results because they are scored low or they don't contain the terms exactly as in the query. This is common when users choose variations of the query terms, including inflections, misspellings, abbreviations, and so on. To improve the user experience, dBlue uses text analysis tools developed by IBM Research to enhance the contents of documents. This process is started by extracting terms from a large collection of documents in the IBM technical support domain to create a domain-specific glossary. The terms in the glossary can consist of canonical form, variant form (inflection, abbreviation, misspelling, etc.), synonym, term definition, statistical data, and other information. This initial glossary is enhanced by eliminating irrelevant terms and reranking terms using special dictionaries and algorithms" (Pages 4-5).

Regarding claim 15, **Doganata** further teaches a computer program comprising: A) a document finder module that comprises the following software sub-modules: a software sub-module that finds documents in the identified categories, using the original textual index (Pages 4-5); and

B) a software sub-module that filters the found documents to find additional relevant documents, based on the identified domain specific terms, synonyms, related terms, and statistically close terms (Pages 4-6).

The examiner notes that **Doganata** teaches **"a document finder module that comprises the following software sub-modules: a software sub-module that finds documents in the identified categories, using the original textual index"** as "The way the system is architected allows combining keyword search with navigational search. Based on a topic or a document type, users can narrow down search findings with a single click. This increases the chances of finding the requested information when the user query isn't specific enough to narrow down the search results on the first attempt. The categorized results are returned with the option of filtering the results based on IBM's product offerings and the document types." (Pages 4-5). The examiner further notes that **Doganata** teaches **"a software sub-module that filters the found documents to find additional relevant documents, based on the identified domain specific terms, synonyms, related terms, and statistically close terms"** as "This increases the chances of finding the requested information when the user query isn't specific enough to narrow down the search results on the first attempt. The categorized results are returned with the option of filtering the results based on IBM's product offerings and the document types" (Pages 4-5), "The search results contain the documents that are indexed against the query terms and scored based on certain statistical criteria. In many real-life situations, the relevant documents can't be found or may not appear at the top of the search results because they are scored low or they don't contain the terms exactly as in the query. This is common when users choose variations of the query terms, including inflections, misspellings, abbreviations, and so on. To improve the user experience, dBlue uses text analysis tools developed by IBM Research to enhance the contents of documents. This process is started by extracting terms from a large collection of documents in the IBM technical support domain to create a domain-specific glossary. The terms in the glossary can consist of canonical form, variant form (inflection, abbreviation, misspelling, etc.), synonym, term definition, statistical data, and other information. This initial glossary is enhanced by eliminating

irrelevant terms and reranking terms using special dictionaries and algorithms" (Pages 4-5), and "In the near future, customers will be able to ask questions in natural language and the system won't require an exact match of words. In the near future, dBlue will also personalize searching so that once a user fills out a profile, responses will be filtered and ranked based on that profile. Multilanguage searches for documents written in Japanese, Chinese, and French will be supported by late 2002" (Page 6).

Regarding claim 16, **Doganata** further teaches a computer program comprising:
A) wherein a meta-data enhancer module comprises the following sub-modules: a software sub-module that creates associations (links) between each found document and the given query (Pages 4-5); and
B) a software sub-module that adds new doc-query links to the meta-data of the corresponding textual index entries (Pages 4-5).

The examiner notes that **Doganata** teaches **"wherein a meta-data enhancer module comprises the following sub-modules: a software sub-module that creates associations (links) between each found document and the given query"** as "The search results contain the documents that are indexed against the query terms and scored based on certain statistical criteria. In many real-life situations, the relevant documents can't be found or may not appear at the top of the search results because they are scored low or they don't contain the terms exactly as in the query. This is common when users choose variations of the query terms, including inflections, misspellings, abbreviations, and so on. To improve the user experience, dBlue uses text analysis tools developed by IBM Research to enhance the contents of documents. This process is started by extracting terms from a large collection of documents in the IBM technical support domain to create a domain-specific glossary. The terms in the glossary can consist of canonical form, variant form (inflection, abbreviation, misspelling, etc.), synonym, term definition, statistical data, and other information. This initial glossary is enhanced by eliminating irrelevant terms and reranking terms using special dictionaries and algorithms. The process of generating and enhancing the glossary is semi-automatic, using glossary tools and the librarian. Figure 4 shows the

multiple components that compose the glossary of technical terms built for the dBlue system" (Pages 4-5) and "Based on the glossary, the important keywords in each document are extracted and ranked, and their related glossary terms (variants, synonyms, etc.) are used to enrich the content of the document. The content enrichment is used to create keyword metatags for biased indexing, expand the query terms to include related terms, and enable search for related documents. To improve the user's search experience, keywords are displayed in the search results and navigating through keywords is possible" (Pages 4-5). The examiner further notes that **Doganata** teaches **"a software sub-module that adds new doc-query links to the meta-data of the corresponding textual index entries"** as "The search results contain the documents that are indexed against the query terms and scored based on certain statistical criteria. In many real-life situations, the relevant documents can't be found or may not appear at the top of the search results because they are scored low or they don't contain the terms exactly as in the query. This is common when users choose variations of the query terms, including inflections, misspellings, abbreviations, and so on. To improve the user experience, dBlue uses text analysis tools developed by IBM Research to enhance the contents of documents. This process is started by extracting terms from a large collection of documents in the IBM technical support domain to create a domain-specific glossary. The terms in the glossary can consist of canonical form, variant form (inflection, abbreviation, misspelling, etc.), synonym, term definition, statistical data, and other information. This initial glossary is enhanced by eliminating irrelevant terms and reranking terms using special dictionaries and algorithms. The process of generating and enhancing the glossary is semi-automatic, using glossary tools and the librarian. Figure 4 shows the multiple components that compose the glossary of technical terms built for the dBlue system" (Pages 4-5) and "Based on the glossary, the important keywords in each document are extracted and ranked, and their related glossary terms (variants, synonyms, etc.) are used to enrich the content of the document. The content enrichment is used to create keyword metatags for biased indexing, expand the query terms to include related terms, and enable search for

related documents. To improve the user's search experience, keywords are displayed in the search results and navigating through keywords is possible" (Pages 4-5).

Regarding claim 17, **Doganata** teaches a search system comprising:

A) a search system analysis system that periodically looks through the search system log and identifies for analysis unsatisfactory <u>customer</u> search queries that do not cite more than a specified number of references (Pages 4-5);

B) a search query analyzer using one or more of the glossary terms, synonyms, known typographical errors and translated words to provide alternative query terms to original <u>query</u> terms in the unsatisfactory <u>customer</u> search queries <u>identified by the search system analysis system</u> (Pages 4-5);

C) a relevant document finder based on enhanced queries including the alternative query terms to locate relevant documents not found by the original <u>unsatisfactory customer search queries identified by the search system analysis system</u> (Pages 4-5); and

D) a meta-data enhancer creating separate enhances links to one or more of said relevant documents linking to said relevant documents the original terms of the unsatisfactory search queries <u>and not found in the relevant documents</u> so that future search queries using the original terms will result in finding said relevant documents not found by the unsatisfactory <u>customer</u> search queries (Pages 4-5).

The examiner notes that **Doganata** teaches **"a search system analysis system that periodically looks through the search system log and identifies for analysis unsatisfactory <u>customer</u> search queries that do not cite more than a specified number of references"** as "The search results contain the documents that are indexed against the query terms and scored based on certain statistical criteria. In many real-life situations, the relevant documents can't be found or may not appear at the top of the search results because they are scored low or they don't contain the terms exactly as in the query. This is common when users choose variations of the query terms, including inflections, misspellings, abbreviations, and so on. To improve the user experience, dBlue uses text analysis tools developed by IBM Research to enhance the contents of

documents. This process is started by extracting terms from a large collection of documents in the IBM technical support domain to create a domain-specific glossary. The terms in the glossary can consist of canonical form, variant form (inflection, abbreviation, misspelling, etc.), synonym, term definition, statistical data, and other information. This initial glossary is enhanced by eliminating irrelevant terms and reranking terms using special dictionaries and algorithms" (Pages 4-5) and "Based on the glossary, the important keywords in each document are extracted and ranked, and their related glossary terms (variants, synonyms, etc.) are used to enrich the content of the document. The content enrichment is used to create keyword metatags for biased indexing, expand the query terms to include related terms, and enable search for related documents. To improve the user's search experience, keywords are displayed in the search results and navigating through keywords is possible" (Pages 4-5). The examiner further notes that **Doganata** teaches **"a search query analyzer using one or more of the glossary terms, synonyms, known typographical errors and translated words to provide alternative query terms to original <u>query</u> terms in the unsatisfactory <u>customer</u> search queries <u>identified by the search system analysis system</u>"** as "The search results contain the documents that are indexed against the query terms and scored based on certain statistical criteria. In many real-life situations, the relevant documents can't be found or may not appear at the top of the search results because they are scored low or they don't contain the terms exactly as in the query. This is common when users choose variations of the query terms, including inflections, misspellings, abbreviations, and so on. To improve the user experience, dBlue uses text analysis tools developed by IBM Research to enhance the contents of documents. This process is started by extracting terms from a large collection of documents in the IBM technical support domain to create a domain-specific glossary. The terms in the glossary can consist of canonical form, variant form (inflection, abbreviation, misspelling, etc.), synonym, term definition, statistical data, and other information. This initial glossary is enhanced by eliminating irrelevant terms and reranking terms using special dictionaries and algorithms" (Pages 4-5) and "Based on the glossary, the important keywords in each document are extracted and ranked, and

their related glossary terms (variants, synonyms, etc.) are used to enrich the content of the document. The content enrichment is used to create keyword metatags for biased indexing, expand the query terms to include related terms, and enable search for related documents. To improve the user's search experience, keywords are displayed in the search results and navigating through keywords is possible" (Pages 4-5). The examiner further notes that **Doganata** teaches **"a relevant document finder based on enhanced queries including the alternative query terms to locate relevant documents not found by the original <u>unsatisfactory customer search queries identified by the search system analysis system</u>"** as "The search results contain the documents that are indexed against the query terms and scored based on certain statistical criteria. In many real-life situations, the relevant documents can't be found or may not appear at the top of the search results because they are scored low or they don't contain the terms exactly as in the query. This is common when users choose variations of the query terms, including inflections, misspellings, abbreviations, and so on. To improve the user experience, dBlue uses text analysis tools developed by IBM Research to enhance the contents of documents. This process is started by extracting terms from a large collection of documents in the IBM technical support domain to create a domain-specific glossary. The terms in the glossary can consist of canonical form, variant form (inflection, abbreviation, misspelling, etc.), synonym, term definition, statistical data, and other information. This initial glossary is enhanced by eliminating irrelevant terms and reranking terms using special dictionaries and algorithms" (Pages 4-5) and "Based on the glossary, the important keywords in each document are extracted and ranked, and their related glossary terms (variants, synonyms, etc.) are used to enrich the content of the document. The content enrichment is used to create keyword metatags for biased indexing, expand the query terms to include related terms, and enable search for related documents. To improve the user's search experience, keywords are displayed in the search results and navigating through keywords is possible" (Pages 4-5). The examiner further notes that **Doganata** teaches **"a meta-data enhancer creating separate enhances links to one or more of said relevant documents linking to said relevant documents the original terms of the**

unsatisfactory search queries <u>and not found in the relevant documents</u> so that future search queries using the original terms will result in finding said relevant documents not found by the unsatisfactory <u>customer</u> search queries" as "The search results contain the documents that are indexed against the query terms and scored based on certain statistical criteria. In many real-life situations, the relevant documents can't be found or may not appear at the top of the search results because they are scored low or they don't contain the terms exactly as in the query. This is common when users choose variations of the query terms, including inflections, misspellings, abbreviations, and so on. To improve the user experience, dBlue uses text analysis tools developed by IBM Research to enhance the contents of documents. This process is started by extracting terms from a large collection of documents in the IBM technical support domain to create a domain-specific glossary. The terms in the glossary can consist of canonical form, variant form (inflection, abbreviation, misspelling, etc.), synonym, term definition, statistical data, and other information. This initial glossary is enhanced by eliminating irrelevant terms and reranking terms using special dictionaries and algorithms" (Pages 4-5) and "Based on the glossary, the important keywords in each document are extracted and ranked, and their related glossary terms (variants, synonyms, etc.) are used to enrich the content of the document. The content enrichment is used to create keyword metatags for biased indexing, expand the query terms to include related terms, and enable search for related documents. To improve the user's search experience, keywords are displayed in the search results and navigating through keywords is possible" (Pages 4-5).

Regarding claim 18, **Doganata** further teaches a search system comprising:
A) wherein said meta-data enhancer links the <u>alternative</u> query terms to the original query terms to <u>automatically</u> locate said relevant documents (Pages 4-5).

The examiner notes that **Doganata** teaches **"wherein said meta-data enhancer links the <u>alternative</u> query terms to the original query terms to <u>automatically</u> locate said relevant documents"** as "The search results contain the documents that are indexed against the query terms and scored based on certain statistical criteria. In

many real-life situations, the relevant documents can't be found or may not appear at the top of the search results because they are scored low or they don't contain the terms exactly as in the query. This is common when users choose variations of the query terms, including inflections, misspellings, abbreviations, and so on. To improve the user experience, dBlue uses text analysis tools developed by IBM Research to enhance the contents of documents. This process is started by extracting terms from a large collection of documents in the IBM technical support domain to create a domain-specific glossary. The terms in the glossary can consist of canonical form, variant form (inflection, abbreviation, misspelling, etc.), synonym, term definition, statistical data, and other information. This initial glossary is enhanced by eliminating irrelevant terms and reranking terms using special dictionaries and algorithms. The process of generating and enhancing the glossary is semi-automatic, using glossary tools and the librarian. Figure 4 shows the multiple components that compose the glossary of technical terms built for the dBlue system" (Pages 4-5) and "Based on the glossary, the important keywords in each document are extracted and ranked, and their related glossary terms (variants, synonyms, etc.) are used to enrich the content of the document. The content enrichment is used to create keyword metatags for biased indexing, expand the query terms to include related terms, and enable search for related documents. To improve the user's search experience, keywords are displayed in the search results and navigating through keywords is possible" (Pages 4-5).

Regarding claim 19, **Doganata** further teaches a search system comprising:
A) wherein the relevant document finder module comprises the following sub-modules: a sub-module that finds documents in identified categories, using the original textual index (Pages 4-5); and
B) a sub-module that filters the found documents to find additional relevant documents, based on the identified domain specific terms, synonyms, related terms, and statistically close terms (Pages 4-6).

The examiner notes that **Doganata** teaches **"wherein the relevant document finder module comprises the following sub-modules: a sub-module that finds**

documents in identified categories, using the original textual index" as "The way the system is architected allows combining keyword search with navigational search. Based on a topic or a document type, users can narrow down search findings with a single click. This increases the chances of finding the requested information when the user query isn't specific enough to narrow down the search results on the first attempt. The categorized results are returned with the option of filtering the results based on IBM's product offerings and the document types." (Pages 4-5). The examiner further notes that **Doganata** teaches **"a sub-module that filters the found documents to find additional relevant documents, based on the identified domain specific terms, synonyms, related terms, and statistically close terms"** as "This increases the chances of finding the requested information when the user query isn't specific enough to narrow down the search results on the first attempt. The categorized results are returned with the option of filtering the results based on IBM's product offerings and the document types" (Pages 4-5), "The search results contain the documents that are indexed against the query terms and scored based on certain statistical criteria. In many real-life situations, the relevant documents can't be found or may not appear at the top of the search results because they are scored low or they don't contain the terms exactly as in the query. This is common when users choose variations of the query terms, including inflections, misspellings, abbreviations, and so on. To improve the user experience, dBlue uses text analysis tools developed by IBM Research to enhance the contents of documents. This process is started by extracting terms from a large collection of documents in the IBM technical support domain to create a domain-specific glossary. The terms in the glossary can consist of canonical form, variant form (inflection, abbreviation, misspelling, etc.), synonym, term definition, statistical data, and other information. This initial glossary is enhanced by eliminating irrelevant terms and reranking terms using special dictionaries and algorithms" (Pages 4-5), and "In the near future, customers will be able to ask questions in natural language and the system won't require an exact match of words. In the near future, dBlue will also personalize searching so that once a user fills out a profile, responses will be filtered and ranked

based on that profile. Multilanguage searches for documents written in Japanese, Chinese, and French will be supported by late 2002" (Page 6).

Regarding claim 20, **Doganata** further teaches a computer program comprising:
A) wherein the meta-data enhancer module comprises the following sub-modules: a sub-module that creates associations (links) between each found document and the given query (Pages 4-5); and
B) a sub-module that adds new doc-query links to the meta-data of the corresponding textual index entries (Pages 4-5).

The examiner notes that **Doganata** teaches **"wherein the meta-data enhancer module comprises the following sub-modules: a sub-module that creates associations (links) between each found document and the given query"** as "The search results contain the documents that are indexed against the query terms and scored based on certain statistical criteria. In many real-life situations, the relevant documents can't be found or may not appear at the top of the search results because they are scored low or they don't contain the terms exactly as in the query. This is common when users choose variations of the query terms, including inflections, misspellings, abbreviations, and so on. To improve the user experience, dBlue uses text analysis tools developed by IBM Research to enhance the contents of documents. This process is started by extracting terms from a large collection of documents in the IBM technical support domain to create a domain-specific glossary. The terms in the glossary can consist of canonical form, variant form (inflection, abbreviation, misspelling, etc.), synonym, term definition, statistical data, and other information. This initial glossary is enhanced by eliminating irrelevant terms and reranking terms using special dictionaries and algorithms. The process of generating and enhancing the glossary is semi-automatic, using glossary tools and the librarian. Figure 4 shows the multiple components that compose the glossary of technical terms built for the dBlue system" (Pages 4-5) and "Based on the glossary, the important keywords in each document are extracted and ranked, and their related glossary terms (variants, synonyms, etc.) are used to enrich the content of the document. The content

enrichment is used to create keyword metatags for biased indexing, expand the query terms to include related terms, and enable search for related documents. To improve the user's search experience, keywords are displayed in the search results and navigating through keywords is possible" (Pages 4-5). The examiner further notes that **Doganata** teaches **"a sub-module that adds new doc-query links to the meta-data of the corresponding textual index entries"** as "The search results contain the documents that are indexed against the query terms and scored based on certain statistical criteria. In many real-life situations, the relevant documents can't be found or may not appear at the top of the search results because they are scored low or they don't contain the terms exactly as in the query. This is common when users choose variations of the query terms, including inflections, misspellings, abbreviations, and so on. To improve the user experience, dBlue uses text analysis tools developed by IBM Research to enhance the contents of documents. This process is started by extracting terms from a large collection of documents in the IBM technical support domain to create a domain-specific glossary. The terms in the glossary can consist of canonical form, variant form (inflection, abbreviation, misspelling, etc.), synonym, term definition, statistical data, and other information. This initial glossary is enhanced by eliminating irrelevant terms and reranking terms using special dictionaries and algorithms. The process of generating and enhancing the glossary is semi-automatic, using glossary tools and the librarian. Figure 4 shows the multiple components that compose the glossary of technical terms built for the dBlue system" (Pages 4-5) and "Based on the glossary, the important keywords in each document are extracted and ranked, and their related glossary terms (variants, synonyms, etc.) are used to enrich the content of the document. The content enrichment is used to create keyword metatags for biased indexing, expand the query terms to include related terms, and enable search for related documents. To improve the user's search experience, keywords are displayed in the search results and navigating through keywords is possible" (Pages 4-5).

10.   Claims 1, 4-9, and 12-20 are rejected under 35 U.S.C. 102(b) based upon a public use or sale of the invention. Specifically, the cited art of **Doganata et al.**

(published on January 01, 2000) details the dBlue system which enhances meta-data of indexed documents by embedding misspellings, synonyms, etc., just as the claimed invention does.

Response to Arguments

11.   Applicant's arguments with respect to claims 1, 4-9, and 12-20 have been considered but are moot in view of the new ground(s) of rejection.

Conclusion

12.   The prior art made of record and not relied upon is considered pertinent to applicant's disclosure.

U.S. PGPUB 2005/0065773 issued to **Huang et al.** on 24 March 2005. The subject matter disclosed therein is pertinent to that of claims 1, 4-9, and 12-20 (e.g., methods the enhance the meta data of indexed documents).

U.S. PGPUB 2004/0254920 issued to **Brill et al.** on 16 December 2004. The subject matter disclosed therein is pertinent to that of claims 1, 4-9, and 12-20 (e.g., methods the enhance the meta data of indexed documents).

U.S. Patent 7,051,023 issued to **Kapur et al.** on 23 May 2006. The subject matter disclosed therein is pertinent to that of claims 1, 4-9, and 12-20 (e.g., methods the enhance the meta data of indexed documents).

U.S. Patent 7,136,845 issued to **Chandrasekar et al.** on 14 November 2006. The subject matter disclosed therein is pertinent to that of claims 1, 4-9, and 12-20 (e.g., methods the enhance the meta data of indexed documents).

U.S. Patent 6,169,986 issued to **Bowman et al.** on 02 January 2001. The subject matter disclosed therein is pertinent to that of claims 1, 4-9, and 12-20 (e.g., methods the enhance the meta data of indexed documents).

U.S. PGPUB 20040249808 issued to **Azzam et al.** on 09 December 2004. The subject matter disclosed therein is pertinent to that of claims 1, 4-9, and 12-20 (e.g., methods the enhance the meta data of indexed documents).

U.S. PGPUB 2005/0055341 issued to **Haahr et al.** on 10 March 2005. The subject matter disclosed therein is pertinent to that of claims 1, 4-9, and 12-20 (e.g., methods the enhance the meta data of indexed documents).

U.S. Patent 6,941,294 issued to **Flank** on 06 September 2005. The subject matter disclosed therein is pertinent to that of claims 1, 4-9, and 12-20 (e.g., methods the enhance the meta data of indexed documents).

U.S. Patent 7,197,508 issued to **Brown** on 27 March 2007. The subject matter disclosed therein is pertinent to that of claims 1, 4-9, and 12-20 (e.g., methods the enhance the meta data of indexed documents).

U.S. Patent 7,127,456 issued to **Brown** on 24 October 2006. The subject matter disclosed therein is pertinent to that of claims 1, 4-9, and 12-20 (e.g., methods the enhance the meta data of indexed documents).

U.S. Patent 6,338,055 issued to **Haggmann** on 08 January 2002. The subject matter disclosed therein is pertinent to that of claims 1, 4-9, and 12-20 (e.g., methods the enhance the meta data of indexed documents).

U.S. PGPUB 2002/0095621 issued to **Lawton et al.** on 18 July 2002. The subject matter disclosed therein is pertinent to that of claims 1, 4-9, and 12-20 (e.g., methods the enhance the meta data of indexed documents).

Article entitled "An Advanced Enterprise Information Search and Delivery System: Fulfilling IBM's one-Web vision" by **Doganata et al.**, dated 14 October 2002. The subject matter disclosed therein is pertinent to that of claims 1, 4-9, and 12-20 (e.g., methods the enhance the meta data of indexed documents).

U.S. Patent 6,772,150 issued to **Whitman et al.** on 03 August 2004. The subject matter disclosed therein is pertinent to that of claims 1, 4-9, and 12-20 (e.g., methods the enhance the meta data of indexed documents).

Contact Information

13. Any inquiry concerning this communication or earlier communications from the examiner should be directed to Mahesh Dwivedi whose telephone number is (571) 272-2731. The examiner can normally be reached on Monday to Friday 8:20 am – 4:40 pm.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Tim Vo can be reached (571) 272-3642. The fax number for the organization where this application or proceeding is assigned is (571) 273-8300.

　　　Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

/Tim T. Vo/
Supervisory Patent Examiner, Art Unit 2168

Mahesh Dwivedi
Patent Examiner
Art Unit 2168

April 06, 2008
/Mahesh H Dwivedi/
Examiner, Art Unit 2168